Leping Li, Ph.D.

Printable PDF version of this page.

Education:

B.S., Medicinal Chemistry	1987	P.R. China
Ph.D., Medicinal Chemistry	1994	University of North Carolina at Chapel Hill

Summary of Research:

In August, 2003, a few months before I moved from a research fellow position to a tenure-track position, I started to explore methods in promoter sequence analysis. Since then, it has evolved into one of my major focuses of research. I am developing and implementing methods for detecting and discovering functional elements such as the *cis*-regulatory motifs in the promoter regions of genes. Finding the individual *cis*-regulatory elements on a gene is an important initial step for understanding its regulation and function. We use paired promoter sequences (human/mouse ortholog) to aid in identifying the "true" *cis*-regulatory motifs, assuming that functional elements such as the *cis*-regulatory motifs are conserved in closely related species.

Towards this goal, first, we created a putative human-mouse gene ortholog promoter sequence database. Second, in order to effectively mine this data set we have developed a sequence alignment algorithm for identifying conserved segments in the paired promoter regions for human and mouse ortholog genes. We assume that transcription factor binding sites are more likely to be present in conserved (i.e., sequence-similar) regions than in non-conserved regions. Third, we have implemented a computational algorithm that can examine the promoter sequences in the data set and scan to identify binding sites for known transcription factors. Finally, we are developing algorithms based on a mathematical approach called the Gibbs sampler to identify common motifs (both known and unknown) that are present in a set of human and mouse promoter sequences.

Besides this new initiative, I am also developing methods for analysis of microarray data and proteomics data. Notably, I have proposed a method called the genetic algorithm/k-nearest-neighbor (GA/KNN) approach. It is a multivariate stochastic search algorithm which selects a subset of genes that can discriminate between different classes of samples, e.g., normal versus tumor tissue, or unexposed versus exposed tissue. This tool has proved able to identify differentially-expressed genes, and, when used in conjunction with clustering methods, to reveal the existence of subcategories that share characteristic patterns of response (e.g., revealing important tumor subtypes).

Source code and documentation for GA/KNN is available by clicking here http://dir.niehs.nih.gov/dirbb/lifiles/softlic.htm.

Recent Publications:

Li, L., Darden, T.A., Weinberg, C.R., Levine, A.J. and Pedersen L.G. Gene assessment and sample classification for gene expression data using a genetic algorithm/k-nearest neighbor method. *Combinatorial Chemistry and High Throughput Screening*, 2001, *4*, 727.

Li, L., Weinberg, C.R., Darden, T.A. and Pedersen L.G. Gene selection for sample classification based on gene expression data: study of sensitivity to choice of parameters of the GA/KNN method. *Bioinformatics*, 2001, *17*, 1131.

Lobenhofer, E.K., Bennett, L., Li, L., Bushel, P. and Afshari, C. Regulation of DNA replication fork genes by 17B-estradiol. *Mol. Endocrinol.*, 2002, *16*, 1215

Heinloth, A.N., Irwin, R.D., Boorman, G.A., Nettesheim, P., Fannin, R.D., Sieber, S.O., Snell, M.L., Tucker, C.J., Li, L., Travlos, G.S., Vansant, G., Blackshear, P.E., Tennant, R.W., Cunningham, M.L. and Paules, R.S. Gene expression profiling of rat livers reveals early indicators of potential adverse effects. *Toxicol. Sci.*, 2004, *80*, 193.

Li, L., Umbach, D.M., Terry, P. and Taylor, J.A. Application of the GA/KNN method to SELDI proteomics data. *Bioinformatics*, 2004, *20*, 1638.

Biostatistics Branch,
Environmental Diseases and Medicine Program,
Division of Intramural Research,
National Institute of Environmental Health Sciences,
National Institutes of Health,
Department of Health and Human Services.

Last updated: 2004/12/7

Please send any comments, corrections, or inquires regarding this page to Bill qb Quattlebaum (quattleb@niehs.nih.gov).